02-02-00

A

JC772 U.S. PTO
02/01/00

Please type a plus sign (+) inside this box → [+]

PTO/SB/05 (4/98)
Approved for use through 09/30/2000. OMB 0651-0032
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 C.F.R. § 1.53(b))

Attorney Docket No.	6843 US 1
First Inventor or Application Identifier	ANIL M. MURCHING et al
Title	A Process to Extract Regions of.......
Express Mail Label No.	EL415462549US

JC625 U.S. PTO
09/496060
02/01/00

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO: Assistant Commissioner for Patents
Box Patent Application
Washington, DC 20231

1. [X] * Fee Transmittal Form (e.g., PTO/SB/17) (Submit an original and a duplicate for fee processing)
2. [X] Specification [Total Pages 13] (preferred arrangement set forth below)
   - Descriptive title of the Invention
   - Cross References to Related Applications
   - Statement Regarding Fed sponsored R & D
   - Reference to Microfiche Appendix
   - Background of the Invention
   - Brief Summary of the Invention
   - Brief Description of the Drawings (if filed)
   - Detailed Description
   - Claim(s)
   - Abstract of the Disclosure
3. [X] Drawing(s) (35 U.S.C. 113) [Total Sheets 3]
4. Oath or Declaration [Total Pages]
   a. [ ] Newly executed (original or copy)
   b. [ ] Copy from a prior application (37 C.F.R. § 1.63(d)) (for continuation/divisional with Box 16 completed)
      i. [ ] DELETION OF INVENTOR(S) Signed statement attached deleting inventor(s) named in the prior application, see 37 C.F.R. §§ 1.63(d)(2) and 1.33(b).

*NOTE FOR ITEMS 1 & 13: IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY FEES, A SMALL ENTITY STATEMENT IS REQUIRED (37 C.F.R. § 1.27), EXCEPT IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C.F.R. § 1.28).

5. [ ] Microfiche Computer Program (Appendix)
6. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)
   a. [ ] Computer Readable Copy
   b. [ ] Paper Copy (identical to computer copy)
   c. [ ] Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

7. [ ] Assignment Papers (cover sheet & document(s))
8. [ ] 37 C.F.R.§3.73(b) Statement (when there is an assignee) [ ] Power of Attorney
9. [ ] English Translation Document (if applicable)
10. [ ] Information Disclosure Statement (IDS)/PTO-1449 [ ] Copies of IDS Citations
11. [ ] Preliminary Amendment
12. [X] Return Receipt Postcard (MPEP 503) (Should be specifically itemized)
13. [ ] * Small Entity Statement(s) (PTO/SB/09-12) [ ] Statement filed in prior application, Status still proper and desired
14. [ ] Certified Copy of Priority Document(s) (if foreign priority is claimed)
15. [X] Other: APPENDIX - 3 Pages

16. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment:
[X] Continuation [ ] Divisional [ ] Continuation-in-part (CIP) of prior application No: 60 / 118,192

Prior application information: Examiner ______ Group / Art Unit: ______

For CONTINUATION or DIVISIONAL APPS only: The entire disclosure of the prior application, from which an oath or declaration is supplied under Box 4b, is considered a part of the disclosure of the accompanying continuation or divisional application and is hereby incorporated by reference. The incorporation can only be relied upon when a portion has been inadvertently omitted from the submitted application parts.

17. CORRESPONDENCE ADDRESS

[ ] Customer Number or Bar Code Label (Insert Customer No. or Attach bar code label here) or [X] Correspondence address below

Name	Peter D. Symes				
	Grass Valley (U.S.) Inc.				
Address	P.O. Box 599000				
	400 Providence Mine Road				
City	Nevada City	State	California	Zip Code	95959-7900
Country	U.S.A.	Telephone	530 478-3437	Fax	

Name (Print/Type)	Francis I. Gray	Registration No. (Attorney/Agent)	27,788
Signature	Francis I. Gray	Date	Feb. 1, 2000

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231.

TITLE OF THE INVENTION

A PROCESS TO EXTRACT REGIONS OF HOMOGENEOUS COLOR IN A DIGITAL PICTURE

CROSS REFERENCE TO RELATED APPLICATIONS

This is a continuation of provisional U.S. Patent Application Serial No. 60/118,192 filed February 1, 1999, now abandoned.

BACKGROUND OF THE INVENTION

The present invention relates to video data processing, and more particularly to a process for extracting regions of homogeneous color in a digital picture.

Extraction of semantically meaningful visual objects from still images and video has enormous applications in video editing, processing, and compression (as in MPEG-4) as well as in search (as in MPEG-7) applications. Extraction of a semantically meaningful object such as a building, a person, a car etc. may be decomposed into extraction of homogeneous regions of the semantic object and performing a "union" of these portions at a later stage. The homogeneity may be in color, texture, or motion. As an example, extraction of a car is considered as extraction of tires, windows and other glass portions, and the body of the car itself.

What is desired is a process that may be used to extract a homogenous color portion of an object.

BRIEF SUMMARY OF THE INVENTION

Accordingly the present invention provides a process for extracting regions of homogeneous color in a digital picture based on a color gradient field with two methods for computing the gradient field -- a weighted Euclidean distance between moment-based feature vectors and a so-called pmf-based distance metric. The digital picture is divided into blocks, and a feature vector is generated for each block as the set of moments for the data in the block. The maximum distance between each block and its nearest neighbors is determined, using either the weighted Euclidean distance metric or the probability mass function-based distance metric, to generate a gradient value for each block. The set of gradient values define the color gradient field. The gradient field is digitized and smoothed, and then segmented into regions of similar color characteristics using a watershed algorithm.

The objects, advantages and other novel features of the present invention are apparent from the following detailed description when read in conjunction with the appended claims and attached drawing.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

Fig. 1 is a block diagram view of an overall process according to the present invention.

Fig. 2 is an illustrative view of an original image.

Fig. 3 is an illustrative view of a segmentation map of the image of Fig. 2 according to a first embodiment of the present invention.

Fig. 4 is an illustrative view of a segmentation map of the image of Fig. 2 according to a second embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The process described here is block-based, i.e. the digital picture is first divided into many non-overlapping rectangular blocks (in general blocks of other shapes and of different sizes, and use of overlapping blocks may be used), and then spatially adjacent blocks that have similar color properties are merged together. This results in the classification of the picture into several spatially contiguous groups of blocks, each group being homogenous in color.

First, segment a digital picture based on a color gradient field, and then use one of two methods for computing that gradient field. The first method makes use of the weighted Euclidean distance between moment-based feature vectors. The second method makes use of the so-called pmf-based distance metric. The overall process is shown in Fig. 1.

The digital input images are assumed to be in YUV format. If the inputs are in a chrominance sub-sampled format such as 4:2:0, 4:1:1 or 4:2:2, the chrominance data is upsampled to generate 4:4:4 material.

Extract one feature vector for each PxQ block of the input picture. There are two stages in the feature vector generation process. In the first stage, transform the data from the original YUV color co-ordinate system into another co-ordinate system known as CIE -- L*a*b* [see *Fundamentals of*

Digital Image Processing, by Anil K. Jain, Prentice-Hall, Section 3.9]. The latter is known to be a perceptually uniform color system, i.e. the Euclidean distance between two points (or colors) in the CIE -- L*a*b* co-ordinate system corresponds to the perceptual difference between the colors.

The next stage in the feature vector generation process is the calculation of the first N moments of the CIE -- L*a*b* data in each block. Thus, each feature vector has 3N components (N moments in *L*, N moments in *a*, and N moments in *b*). (See the Appendix)

The next stage in the region extraction process is that of gradient extraction. Estimate a block-based gradient field for the input picture (i.e. get one scalar gradient value for each PxQ block of the input picture). The gradient at the (*i, j*)-th block of the input picture is defined as the maximum of the distances between the block's feature vector *f*(i,j) and its nearest neighbor's feature vectors. (See Appendix)

(In the maximization, let k and l each vary from −1 to +1, but do not allow $k = l = 0$ simultaneously! Also, along the borders of the image, consider only those neighboring blocks that lie inside the image boundaries). Use one of two types of distance functions.

Other methods to select the gradient value from the above set of distances, for example the minimum, median, etc. May be used. It is necessary to evaluate the performance of the segmentation algorithm when such methods are used.

The distance function is simply the weighted Euclidean distance between two vectors. (See Appendix). In the formula, the weighting factors may be used to account for the differences in scale among the various moments. This metric is very easy to implement. In one implementation, set N = 1, i.e. use only the mean values within each PxQ block, and set the weighting factors to unity (this makes sense, since the CIE -- L*a*b* space is perceptually uniform).

The second choice of the distance metric is a little more involved. Here, the fact is exploited that using the moments of the data within the PxQ block, an approximation to the probability mass function (pmf) of that data may be computed. The pmf essentially describes the distribution of the data to be composed of a mixture of several values, with respective probabilities. The values and the probabilities together constitute the pmf. Compute these values using the moments as described in the Appendix.

Thus, the moment-based feature vector of each PxQ block may be converted into a pmf-based representation. With such a representation, then the distance between two feature vectors may be computed via the distance between the two pmf's. For this, make use of the Kolmogorov-Smirnoff (K-S) test, as described in Section 14.3 of *"Numerical Recipes in C"*, 2nd edition, by W. A. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Cambridge University Press. (Essentially, the distance between two pmf's is the area under the absolute value of the difference between

the two cumulative distribution functions, see the above-mentioned chapter for details).

Though the K-S test is prescribed for pmf's of a single variable, the data is in fact three-dimensional (*L*, *a*, and *b* components). Strictly speaking, it is necessary to compute the joint, three-dimensional pmf, and then compute a distance between two pmf's. This is however a very hard problem to solve, and instead a simplifying assumption is made. Assume that the color data in a PxQ block may be modeled by means of three independent pmf's, one each for the *L*, *a*, and *b* components. (See Appendix)

The gradient field, as computed above, yields values that lie along the positive real axis (i.e. can vary from zero to infinity). In practice, the gradient values occupy a finite range, say from minimum to maximum. Digitize the gradient field at a precision of B bits, by dividing the above range into 2^B levels. In one implementation, choose B = 8.

After the gradient field has been digitized, perform morphological preprocessing. This process removes small bumps in the gradient field, and helps the subsequent watershed algorithm to perform a better segmentation. The preprocessing algorithm used has been taken from "*Unsupervised Video Segmentation Based on Watersheds and Temporal Tracking*", by Demin Wang, pages 539 through 546, IEEE Transactions on Circuits and Systems for Video Technology, Volume 8, Number 5, September 1998. "Reconstruction By Erosion" is used as described in

"*Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms*", by Luc Vincent, pages 176 through 201, IEEE Transactions on Image Processing, Volume 2, Issue 2, April 1993. In this process, a smoothing threshold that is 0.7% of the dynamic range of the gradient field is used.

The digitized gradient field, after the above preprocessing, is segmented by what is known as the watershed algorithm. The algorithm description is in the above-mentioned journal article by Luc Vincent. The watershed algorithm divides the gradient field into a set of spatially connected regions, each of which is "smooth" in its interior. Thus, these regions are characterized by having strong gradients at their boundaries. Since the gradient value is proportional to the perceptual difference in color, by the above way of calculating the distance metric, the image is segmented into regions of homogenous color.

Once the input digital image has been segmented into regions that are homogenous in color and are spatially connected, this information may be used in database/search applications. Each region may be represented by one feature vector, consisting of either the same N moments that were used in the segmentation process, or consisting of the pmf-based representation that are computed from those moments. The latter representation is more powerful, because capturing the probability distribution of the data is known to be very useful for indexing visual objects for search applications. In this case the work by Szego

("Orthogonal Polynomials", 4th edition, American Math. Society, Providence, Volume 23, 1975) is used to compute the pmf-based representation from the moments. Then, create an entry for this image in the database, consisting of the classification map together with the characteristic feature vector for each class (region). The use of such an index for database applications is described in a co-pending provisional U.S. Patent Application Serial No.60/118,.

Although in the described implementation non-overlapping rectangular blocks are used, this process may be generalized to blocks of other shapes (square, hexagonal, etc.). Also overlapping blocks may be used, which helps in obtaining a segmentation map that is of higher resolution (than the current block-based segmentation map).

One particular computation of local activity measures has been described, where the moments are computed over rectangular (PxQ) blocks. Activity measures other than moments may be used. Also different block sizes for different areas of the image may be used.

The described pmf-based distance metric uses only two representative values and their probabilities. This metric may be extended by using more representative values (resulting in a more accurate representation of the true probability distribution of the data). A closed form solution for computing more representative values and their corresponding probabilities can be found in the work by Szego.

Other methods than the watershed algorithm may be used to merge blocks. K-means clustering, quadtree segmentation, etc. are possible alternatives.

Thus the present invention provides a process for extracting regions of homogeneous color in a digital picture by segmenting the picture based on a color gradient field, computing the gradient field by one of two distance metrics, digitizing and preprocessing the gradient field, and then segmenting the preprocessed digitized color gradient field with a watershed algorithm.

CLAIM OR CLAIMS

WHAT IS CLAIMED IS:

1. A method of extracting regions of homogeneous color in a digital picture comprising the steps of:

dividing the digital picture into blocks; and

merging together spatially adjacent blocks that have similar color properties to extract the regions of homogeneous color.

2. The method as recited in claim 1 wherein the merging step comprises the steps of:

extracting a feature vector for each block;

estimate a scalar gradient value for each block as a function of the feature vector, the set of gradient values defining a color gradient field;

digitizing the color gradient field;

preprocessing the digitized color gradient field to produce a smoothed color gradient field; and

segmenting the smoothed color gradient field with a watershed algorithm that divides the smoothed color gradient field into a set of spatially connected regions of homogeneous color.

3. The method as recited in claim 2 wherein the extracting step comprises the steps of:

transforming data in each block into a perceptually uniform color system; and

calculate N moments of the data in each block for each color component, the set of moments being the feature vector for the block.

4. The method as recited in claim 2 wherein the estimating step comprises the steps of:

obtaining distances between the feature vector of each block and the feature vectors of each neighboring block; and

selecting the maximum of the distances as the gradient value for the block.

5. The method as recited in claim 4 wherein the obtaining step comprises the steps of:

applying a weighted Euclidean distance metric to the feature vectors to obtain the distances.

6. The method as recited in claim 4 wherein the obtaining step comprises the steps of:

converting the feature vector of each block into a probability mass function-based representation for each color component;

computing distances between the probability mass function-based representations of each block and the corresponding probability mass

function-based representations of each neighboring block; and

selecting the maximum distance of the probability mass function-based representations as the gradient value for the block.

ABSTRACT OF THE DISCLOSURE

A method of extracting regions of homogeneous color from a digital picture divides the digital picture into blocks and generates a feature vector for each block as a set of moments of the data for the block. The distance between the feature vector of each block and the feature vectors of the nearest neighboring blocks are determined using either a weighted Euclidean distance metric or a probability mass function-based distance metric. The maximum distance is the gradient value for the block, and the set of gradient values over all the blocks form a color gradient field. The gradient field is digitized and smoothed, and then segmented into regions of similar color characteristics using a watershed algorithm.

2 Overall Process

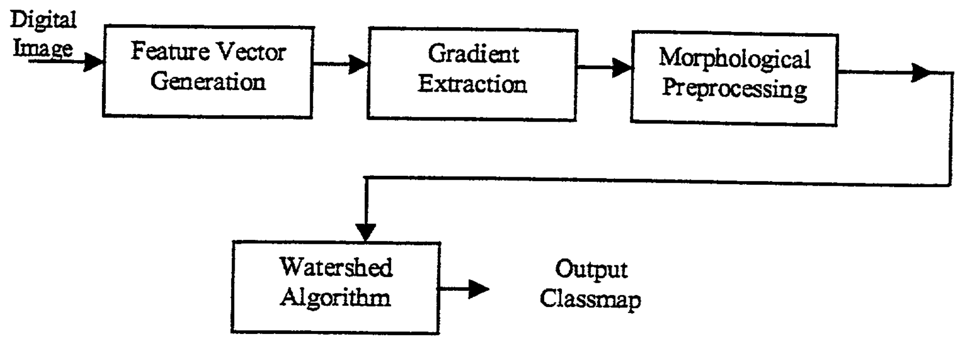

Figure 1 Overall process used in region extraction

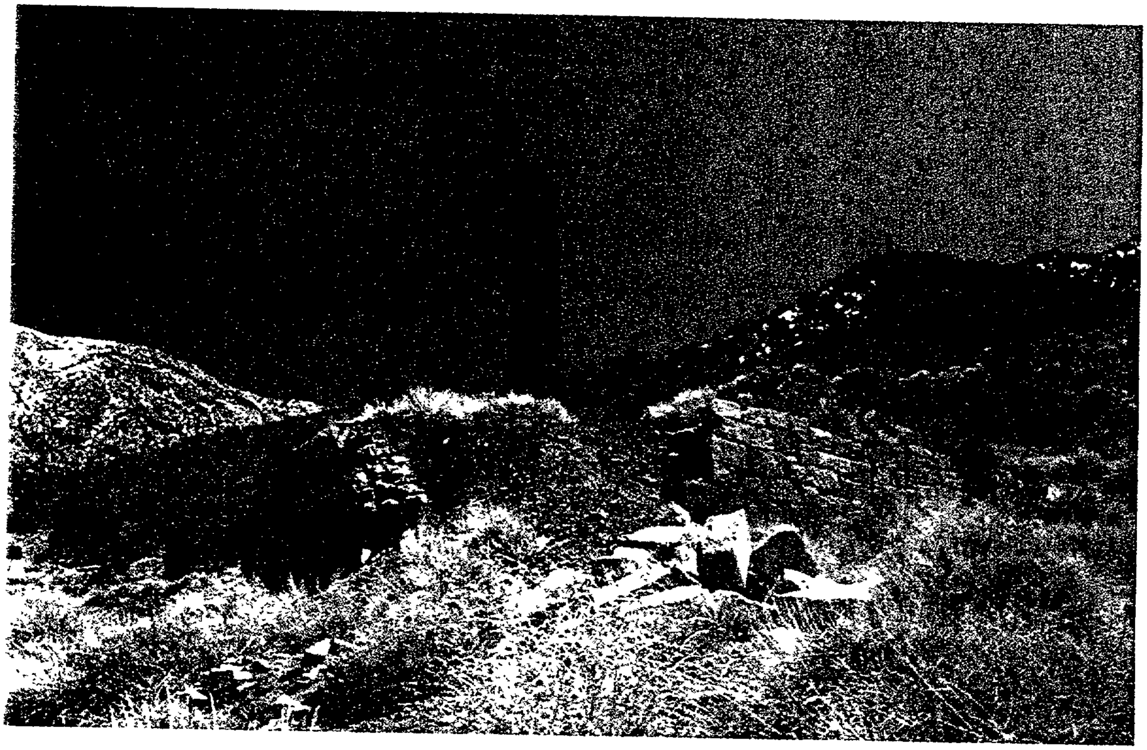

Figure 2 Original image

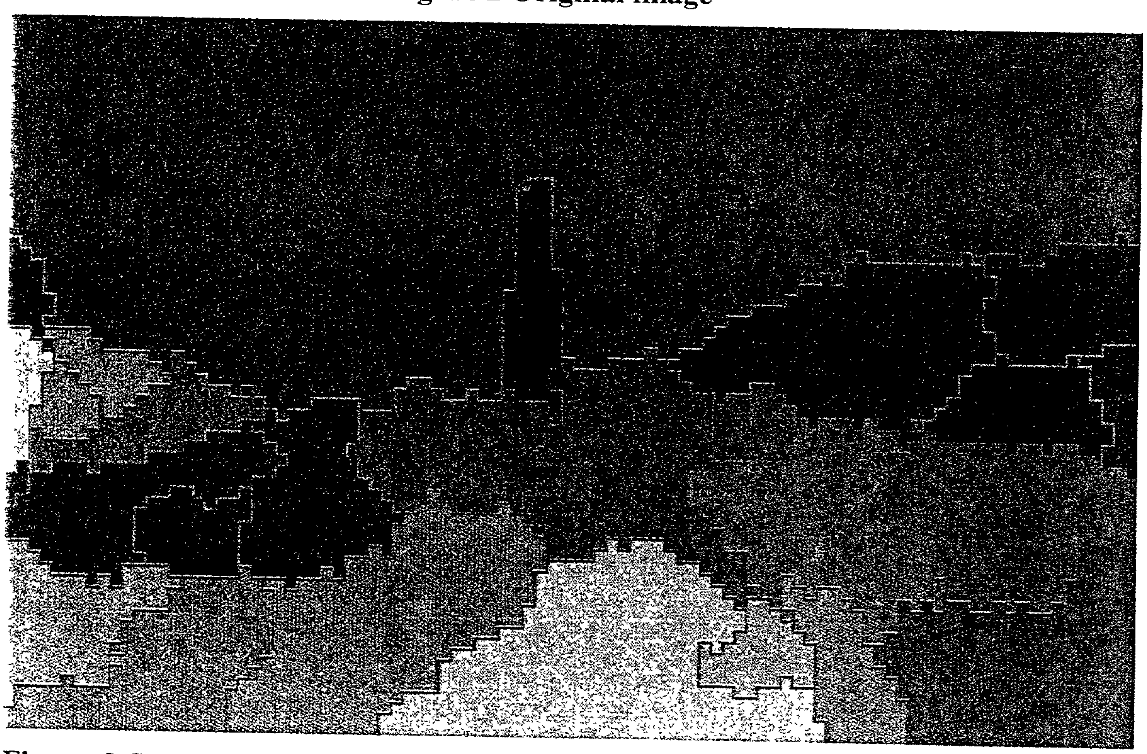

Figure 3 Segmentation map obtained by merging 8x8 blocks, and using pmf-based distance metric (a total of N = 9 moments are computed for each block)

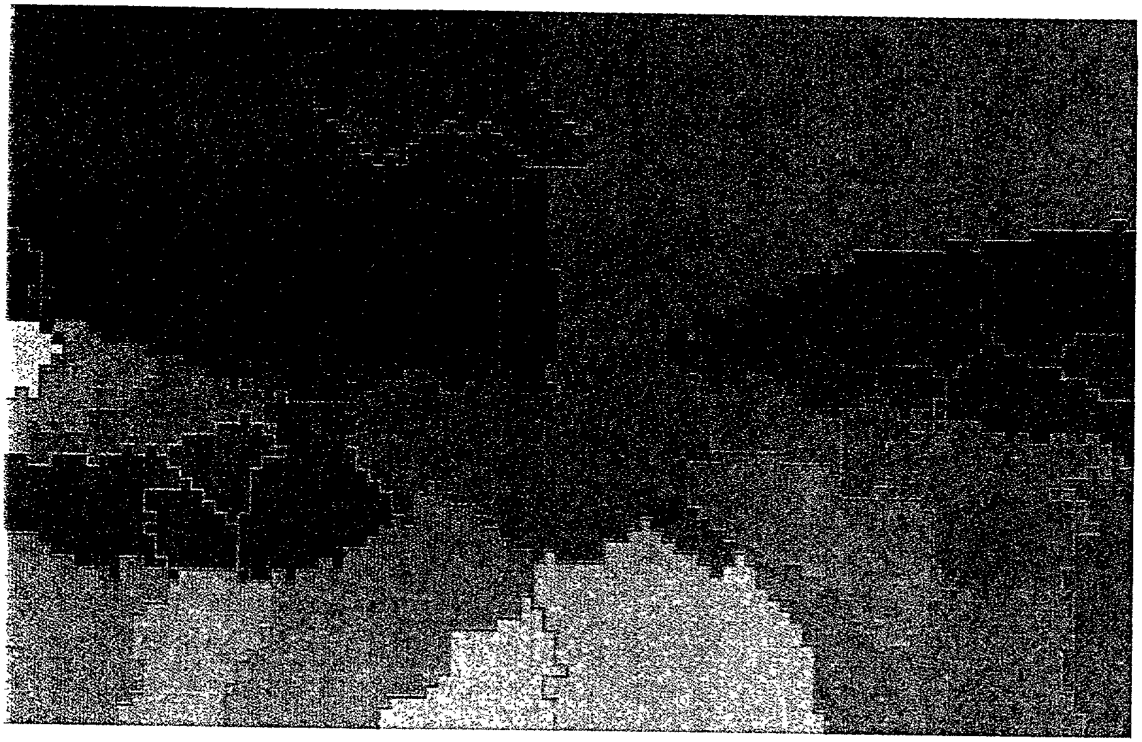

Figure 4 Segmentation map obtained by merging 8x8 blocks, using Euclidean distance metric (only 3 moments, namely the mean values, are computed in each block)

2 Overall Process

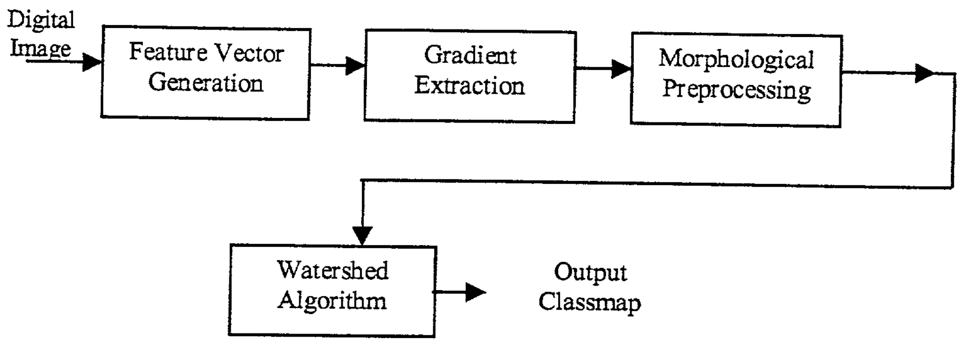

Figure 1 Overall process used in region extraction

3 Input Image Data

The digital input images are assumed to be in YUV format. If the inputs are in a chrominance sub-sampled format such as 420, 411 or 422, the chrominance data is upsampled to generate 444 material.

4 Feature Vector Generation

We extract one feature vector for each PxQ block of the input picture. There are two stages in the feature vector generation process. In the first stage, we transform the data from the original YUV color co-ordinate system into another co-ordinate system known as $CIE-L^*a^*b^*$ [see *Fundamentals of Digital Image Processing*, by Anil K. Jain, Prentice-Hall, Section 3.9]. The latter is known to be a perceptually uniform color system, i.e. the Euclidean distance between two points (or colors) in the $CIE-L^*a^*b^*$ co-ordinate system corresponds to the perceptual difference between the colors.

The next stage in the feature vector generation process is the calculation of the first N moments of the $CIE-L^*a^*b^*$ data in each block. Thus, each feature vector has 3N components (N moments in L, N moments in a, and N moments in b). We can denote the (3Nx1) feature vector of the (i, j)-th block of the input picture as follows.

$$\vec{f}(i,j) = [{}_L m_1, \ldots, {}_L m_N, {}_a m_1, \ldots, {}_a m_N, {}_b m_1, \ldots, {}_b m_N]^T,$$

where the k-th moment in, say, the L component, is given by

$${}_L m_k = \frac{1}{PQ} \sum L^k(x,y),$$

where (x, y) represents the index of a point in the (i, j)-th block.

5 Gradient Extraction

The next stage in our region extraction process is that of gradient extraction. We will estimate a block-based gradient field for the input picture (i.e. we get one scalar gradient value for each PxQ block of the input picture). The gradient at the *(i, j)*-th block of the input picture is defined as the maximum of the distances between the block's feature vector $\vec{f}(i,j)$ and its nearest neighbor's feature vectors.

$$grad(i,j) = \max_{k,l\in\{-1,0,1\}} \{d[\vec{f}(i,j), \vec{f}(i-k, j-l)]\},$$

where d[.,.] is function that assigns a distance value to a pair of feature vectors. (Note: in the above maximization, we let k and l each vary from -1 to $+1$, but do not allow $k = l = 0$ simultaneously! Also, along the borders of the image, we consider only those neighboring blocks that lie inside the image boundaries). In our work, we will employ two types of distance functions.

We could use other methods to select the gradient value from the above set of distances, for example the minimum, median, etc. We need to evaluate the performance of the segmentation algorithm when such methods are used.

5.1 Weighted Euclidean Distance Metric

Here, the distance function d[.,.] is simply the weighted Euclidean distance between the two vectors.

$$d[\vec{f}(i,j), \vec{g}(k,l)] = \sqrt{\begin{array}{l} \{{}_L w_1 ({}_L m_{1,f} - {}_L m_{1,g})^2 + \ldots + {}_L w_N ({}_L m_{N,f} - {}_L m_{N,g})^2 + \\ {}_a w_1 ({}_a m_{1,f} - {}_a m_{1,g})^2 + \ldots + {}_a w_N ({}_a m_{N,f} - {}_a m_{N,g})^2 + \\ {}_b w_1 ({}_b m_{1,f} - {}_b m_{1,g})^2 + \ldots + {}_b w_N ({}_b m_{N,f} - {}_b m_{N,g})^2\} \end{array}}, \text{ where}$$

$$\vec{g}(k,l) \equiv \vec{f}(i = k, j - l).$$

In the above formula, the weighting factors $\{{}_{\{\}} w_{\{\}}\}$ can be used to account for the differences in scale among the various moments. This metric is very easy to implement. In our implementation, we set N = 1, i.e. use only the mean values within each PxQ block, and set the weighting factors to unity (this makes sense, since the $CIE - L^* a^* b^*$ space is perceptually uniform).

5.2 Probability Mass Function Based Distance Metric

The second choice of the distance metric is a little more involved. Here, we exploit the fact that using the moments of the data within the PxQ block, we can compute an approximation to the probability mass function (pmf) of that data. The pmf essentially describes the distribution of the data to be composed of a mixture of several values $v_0, v_1, v_2, \ldots$, with respective probabilities $P_0, P_1, P_2, \ldots$ The values and the probabilities together constitute the pmf. We can compute these values using the moments as follows. For ease of notation, we will drop the subscripts *L, a,* and *b,* because the equations that we provide apply to all three color components.

Initially, we approximate the distribution as a mixture of two values, v_0 and v_1, with probabilities P_0 and P_1 respectively. We use the moments-based approach given in Ali Tabatabai's Ph.D. thesis to estimate the values v_0, v_1, P_0 and P_1. In this method, we need the first three moments of the data (i.e. N = 3):

$$m_1 = \frac{1}{PQ}\sum L(x,y),$$

$$m_2 = \frac{1}{PQ}\sum L^2(x,y), \text{ and}$$

$$m_3 = \frac{1}{PQ}\sum L^3(x,y),$$

where $L(x,y)$ are data values in the (i,j)-th block. Then,

$$P_0 = \frac{1}{2}\left(1 + s\sqrt{\frac{1}{4+s^2}}\right),$$

$$P_1 = 1 - P_0,$$

$$v_0 = m_1 - \sigma\sqrt{P_1/P_0}, \text{ and}$$

$$v_1 = m_1 + \sigma\sqrt{P_0/P_1}, \text{ where}$$

$$s = \frac{m_3 + 2m_1^3 - 3m_1 m_2}{\sigma^3}, \text{ and}$$

$$\sigma = \sqrt{\{m_2 - (m_1)^2\}}.$$

Thus, we can convert the moment-based feature vector of each PxQ block into a pmf-based representation. Once we have such a representation, then the distance between two feature vectors can be computed via the distance between the two pmf's. For this, we make use of the Kolmogorov-Smirnoff (K-S) test, as described in Section 14.3 of *"Numerical Recipes in C"*, 2nd edition, by W. A. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Cambridge University Press. (Essentially, the distance between two pmf's is the area under the absolute value of the difference between the two cumulative distribution functions, see the above-mentioned chapter for details).

Though the K-S test is prescribed for pmf's of a single variable, the data we have is in fact three-dimensional (*L, a,* and *b* components). Strictly speaking, we need to compute the joint, three-dimensional pmf, and then compute a distance between two pmf's. This is however a very hard problem to solve, and instead, we make a simplifying assumption. We assume that the color data in a PxQ block can be modeled by means of three independent pmf's, one each for the *L, a,* and *b* components. Let us denote these pmf's by pmf_L, pmf_a, and pmf_b respectively. Also, denote the K-S distance measure between two pmfs by $d_{KS}(.,.)$, then, the overall distance metric is given by

$$d[\vec{f}(i,j), \vec{g}(k,l)] = d_{KS}(pmf_{L,f}, pmf_{L,g}) + d_{KS}(pmf_{a,f}, pmf_{a,g}) + d_{KS}(pmf_{b,f}, pmf_{b,g}).$$